Project Abstract

An Exploratory Space-Time Data Analysis Toolkit for Spatial Social Science Research
BCS-0433132
Sergio J. Rey
San Diego State University

Objectives

The study of regional economic growth, inequality, divergence and convergence attracts considerable interest across multiple social sciences. By definition, these analyses rely on data that are spatially referenced. Only a few very recent studies, however, have given attention to the role of spatial dependence and spatial heterogeneity in the empirical analysis of regional economic evolutions. Research in the fields of geographical information systems (GIS), spatial statistics, and spatial econometrics has generated new methods designed to treat these spatial effects, but these methods do not address the dynamic dimensions of regional economic change. A truly integrated social science requires a toolkit that integrates both the spatial and temporal dimensions of socioeconomic phenomena. The objectives of this research project are to develop such a toolkit by (1) examining the implications of spatial clustering and spatial heterogeneity for the application of exploratory data analysis (EDA) techniques in a dynamic context; (2) developing new statistical methods for exploratory space-time data analysis (ESTDA); and (3) implementing these methods in an Open Source package for exploratory space-time analysis of social processes. The methods to be used include exploratory spatial data analysis (ESDA), exploratory temporal data analysis (ETDA), Monte Carlo simulation studies of the empirical properties of the new ESTDA methods, object-oriented programming, and dynamic geovisualization.

The toolkit Space-Time Analysis of Regional Systems (STARS) is an open source package designed for the dynamic exploratory analysis of data measured for areal units at multiple points in time. STARS consists of four core analytical modules: [1] ESDA: exploratory spatial data analysis; [2] Inequality measures; [3] Mobility metrics; [4] Spatial Markov. Developed using the Python object oriented scripting language, STARS lends itself to three main modes of use. Within the context of a command line interface (CLI), STARS can be treated as a package which can be called from within customized scripts for batch oriented analyses and simulation. Alternatively, a graphical user interface (GUI) integrates most of the analytical modules with a series of dynamic graphical views containing brushing and linking functionality to support the interactive exploration of the spatial, temporal and distributional dimensions of socioeconomic and physical processes. Finally, the GUI and CLI modes can be combined for use from the Python shell to facilitate interactive programming and access to the many libraries contained within Python.

Activities and Accomplishments

Research on the project has focused on four areas of activies. The first concerns a set of methodological advances developing new statistical methods for exploratory space-time data

analysis. These include the development of two new tests for spatial autocorrelation in dynamic series (Rey, Janikas, and Smirnov, 2005 and Rey and Janikas 2005b), and a spatially explicit approach to analyzing the evolution of income gaps between economies in a regional system – so called spatial sigma convergence (Rey and Dev, 2005).

The second component of the research effort has been to continue to enhance the STARS package in a number of dimensions. First, we have refactored the data import component of STARS by removing a dependence on shapelib (a c-library for reading shapefiles). As a result a single version of STARS runs on all supported platforms (Windows, Mac OS X, Linux). This component has also been redesigned to improve the "user-friendliness" of data import. The user interface has also been enhanced to incorporate context sensitive help dialogs guiding the user through the options of each analytical method, as well as incorporation of example data sets, graphical tutorials (QuickTime Movies) and a user guide.

In addition to the enhancements of the user interface, STARS has also been extended to include the new analytical methods described above. A key component of this effort has been the development of novel methods for geovisualization. An example screenshot is shown in Figure 1. The user has selected the state of Illinois on the map via a mouse click (lower left). This highlights the associated observations in the three remaining views, each giving a different spatial or temporal perspective. In the upper right view the Moran Scatter Plot depicts each state's income against that of its geographical neighbors. Illinois is seen to have above average income, yet its neighboring states have incomes below the national average.

In the lower right panel, a Conditional Scatter Plot shows that the dynamic relationship between incomes and Illinois and its neighboring states (red line) is markedly different from the overall trend (black line). Finally, the TimePath view in the upper left portrays the co-movement of income for Illinois and its regional neighbors. While Illinois initially had an income substantially above the national average in 1929, it has moved towards the national average over the century. The neighboring states, however, have tended to remain in the same relative positions in the income distribution. The TimePath itself was dynamically created by the user issuing a control-click on the Illinois point in the Moran Scatter Plot.

The third area of activity has been the application of the toolkit to substantive research problems in a number of social sciences including regional economics (Yamamoto, 2005; Janikas and Rey, 2005; Rey and Janikas, 2005a), and spatial epidemiology (Getis, 2005; Anselin, et al 2005).

The final area of activity has been outreach, first through interaction with ESRI to explore potential transfer of the new methods to commercial GIS software (Rey, 2004) and second with other open source spatial analysis projects. The latter has resulted in the establishment of a collaborative project (Anselin and Rey, 2005) to develop common spatial analysis functions in a shared library for use by STARS and other projects (PySpace, GeoVista) as described in (Anselin, et al 2005)

Broader Impacts

This project will make significant contributions to the practice of spatial social science, the modeling of human dynamics, and to basic understanding of the nature of regional growth and inequality. The incorporation of space and time into models of regional inequality and growth will provide more comprehensive and accurate descriptions of human social and economic behavior. The project will develop an exploratory space-time toolkit that will be Open Source and accessible to a broad array of social science researchers to enhance the analysis of human and social dynamics. As such, this research is expected to have implications in areas such as studies of urban segregation patterns, space-time epidemiology and public health, criminology, housing market dynamics, socioeconomic inequalities, among others. From a policy perspective, the development of new spatially explicit measures will provide planners and analysts with capabilities to design policy interventions targeted at key individual geographical areas. By taking spatial spillovers into account, this spatially focused strategy will leverage the impact of such policy programs across the boundaries of a single area, thereby increasing the effectiveness of the policy.

Publications and Presentation

Anselin, L. and S.J. Rey (2005)``PySAL a Python library for spatial analytical functions." *Geocomputation 2005*. University of Michigan. August.

Anselin, L., Rey, S.J., Gahegan, M., and Hardisty, F. (2005) “Prostate cancer, ESDA, and spatial statistics.” Centers for Disease Control. Award numberTS-1125

Bonet, J. (2005) “Inequidad espacial en la dotación educativa regional en Colombia”. Banco de la República, Colombia.

Getis, A. (2005) “Dengue Transmission in Time and Space: Models and Methods.” *GEOMED 2005 is the fifth international, interdisciplinary conference on geomedical systems*. Cambridge. September.

Janikas , M.V. and S.J. Rey (2005) ``Spatial clustering, inequality and income convergence." *Region et Developpement*. Forthcoming.

Rey, S.J. (2004) ``STARS: Space-Time Analysis of Regional Systems." ESRI-University of Redlands Colloquia. September

Rey, S.J. (2005) ``STARS: Space-Time Analysis of Regional Systems." *Workshop on Spatial Econometrics*. Kiel Institute for World Economics, Kiel Germany. April.

Rey, S.J. and L. Anselin (2005) ``Recent Advances in Software for Spatial Analysis in the Social Sciences." *Geographical Analysis*. Forthcoming.

Rey, S.J. and B. Dev (2005)``Sigma-convergence in the presence of spatial effects.". Submitted *Papers in Regional Science*.

Rey, S.J. and M.V. Janikas (2005a) ``Regional convergence, inequality and space." *Journal of Economic Geography*. 5: 155-176.

Rey, S.J. and M.V. Janikas (2005b) ``STARS: Space-Time Analysis of Regional Systems." *Geographical Analysis*. Forthcoming.

Rey, S.J., M.V. Janikas and Smirnov, O. (2005) ``Exploratory geovisualization of spatial dynamics." *Geocomputation 2005*. University of Michigan. August.

Yamamoto, D. (2005) "Evolution of Regional Per Capita Income Differentials in the United States and Japan: A Comparative Study." *Association of American Geographers*, Denver.

Project Website

http://stars-py.sf.net/

Figure 1. Example multidimensional views of spatial dynamics.